Pattern recognition in trading behaviors before stock price jumps: new method based on multivariate time series classification

Ao Kong*

School of Finance, Nanjing University of Finance and Economics, Nanjing 210023, China

Robert Azencott

Department of Mathematics, University of Houston, TX 77204, USA

Hongliang Zhu

School of Management and Engineering, Nanjing University, Nanjing 210023, China

Abstract

This paper extends the work of Boudt and Pertitjean (2014) and investigates the trading patterns before price jumps in the stock market based on a new multivariate time classification technique. Different from Boudt and Pertitjean (2014), our analyzing scheme can explore the "time-series information" embedded in the trading-related attributes and provides a set of jump indicators for abnormal pattern recognition. In addition to the commonly used liquidity measures, our analysis also involves a set of technical indicators to describe the micro-trading behaviors. An empirical study is conducted on the level-2 data of the constituent stocks of China Security Index 300. It is found that among all the candidate attributes, several volume and volatility-related attributes exhibit the most significant abnormality before price jumps. Though some of the abnormalities start just shortly before the occurrence of the jumps, some start much earlier. We also find that most of our attributes have low mutual dependencies with each other from the perspective of time-series analysis, which allows various perspectives to study the market trading behaviors. To this end, our experiment provides a set of jump indicators that can effectively detect the stocks with extremely abnormal trading behaviors before price jumps. More importantly, our study offers a new framework and potential useful directions for trading-related pattern recognition problem using the time series classification techniques.

Keywords: price jumps, high-frequency data, mutual information, multivariate time series analysis, pattern recognition

1. Introduction

Large and discontinuous changes, known as jumps, are essential components of stock price dynamics(Merton, 1976). With the perceived risk differing from small and regular price movements, they play an important role in the risk measurement, option pricing, and portfolio allocation(Duffie

and Pan, 2001; Jarrow and Rosenfeld, 1984; Kapadia and Zekhnini, 2019; Zhou et al., 2019).As a result, investigating the trading patterns before price jumps, as well as exploring the jump indicators in the stock market are of great importance in the financial field (Farmer et al., 2004; Lee and Mykland, 2008; Boudt and Pertitjean, 2014). Even so, limited literature can be found related to this topic. It is possible because the jumps, occurring instantly, often follow micro precursors hiding in the complex noise of price dynamics.

Jiang and Lo (2011) was the first to study the liquidity dynamics before price jumps, but based on the treasury market and concentrate on the effect of microeconomic news announcements. Regarding the stock market, Boudt and Pertitjean (2014) was the first to investigate the characteristics of tick data before the jumps in the U.S. market based on more than ten liquidity measures and found that some measures, such as the trading volume, the number of trades, the quoted and effective price spread exhibit significant abnormal movements. Following the work of Boudt and Pertitjean (2014), Będowska-Sójka (2016) and Wan et al. (2017) then revealed some different liquidity dynamics before the price jumps in the Polish and Chinese stock markets based on tick and level-2 data respectively. Despite the different markets, the statistical methods in the these studies are similar: after extracting the values of the liquidity measure within a prior window before intraday price jumps, a Mann-Whitney test is used to compare their median values to those within non-jumping days from time to time. It is worthy to note that the common Mann-Whitney test treats the liquidity series as a set of independent random variables, thus has neglected the information embedded in the whole series.

Our work is then an extension of the work of Boudt and Pertitjean (2014), as well as the following works related to liquidity pattern recognition in the stock market(Wan et al., 2017; Będowska-Sójka, 2016). Different from them, we investigate the liquidity dynamics based on a new technique for multivariate time series classification(Ircio et al., 2020). This technique takes into account the information and properties of the time series and provides a tool to select a highly informative combination of attributes for pattern recognition. Among all the previous methods for multivariate time series analysis (Fang et al., 2015; Jovic and Jovic, 2017; Han et al., 2015; Yoon et al., 2005; He et al., 2019), we choose this technique because it does not transform the original time series into a different representation; thus, the issue of losing interpretability of selected features has been avoided. Apart from the liquidity measures, we also involve a number of technical indicators in our study to describe the intraday trading behaviors, the idea of which is inspired by our previous work (Kong et al., 2020). Time series-based mutual information, as a key point of the classification technique, is used to evaluate the abnormality embedded in the variables related to the occurrence of price jumps, which is more suitable to analyze the non-linear dependency between random variables with high noise levels while do not need to assume their distributions(Shannon, 2001). The goal of our study is to select a generic set of jump indicators

that can represent the abnormal patterns of trading behaviors before price jumps for all stocks, and use them to identify the common and idiosyncratic patterns of the stocks in the market. On top of these, our study aims to provide a new framework for this trading-related pattern recognition problem using time series analyzing tools.

Our paper is also related to several papers that analyse the liquidity dynamics before firm-specific news(Ranaldo, 2008; Lakhal, 2008; Riordan et al., 2013). Although firm-specific news can force stock price jumps, conditioning the intraday analysis on these scenarios does not completely capture the trading behaviors of market participants before stock price jumps since many jumps are not directly associated with public news(Joulin et al., 2008; Lahaye et al., 2011).

The main contribution of our paper are in several aspects. First, we propose a systematic framework based on a new multivariate time series classification technique to study the liquidity dynamics before stock price jumps. Compared to existing method, our approach takes into account the trading information along the time scale when evaluating the information level of trading-related attributes, so that the duration and the level of the abnormality can be better explored and compared among attributes; on the other hand, our approach can evaluate the combinatory power of the candidate attributes in recognizing the abnormal patterns as well as select a set of jump indicators for trading-related pattern analysis of individual stocks. Second, in addition to the commonly used liquidity measures, our experiment incorporates a set of technical indicators, which have low mutual dependencies with each other and with the liquidity measures from the perspective of time-series information. Then the whole set of candidate attributes allow to assess the trading behaviors from various perspectives. Third, based on our approach, we find that although most of the abnormalities are present just shortly before the occurrence of the jumps, some start much earlier. This complement the literature on the study of starting point of the abnormal trading movements from the perspective of time series analysis. Forth, our study provides a generic set of jump indicators for abnormal pattern recognition, which can effectively detect the stocks that have extremely abnormal trading behaviors before price jumps.

The remainder of this paper is organized as follows. Section 2 provides the preliminaries of the new multivariate time series classification technique. Section 3 details our proposed experimental methodology. Section 4 shows the experimental results while section 5 concludes.

2. Preliminaries: Mutual information-based multivariate time series analysis

Consider P random variables $TS^{(1)}, TS^{(2)}, \cdots, TS^{(P)}$, whose realizations are time series. Assume we have a set of samples $\{X_1, X_2, \cdots, X_N\}$, where $X_n = (ts_n^{(1)}, ts_n^{(2)}, \cdots, ts_n^{(P)})$, $n \in \{1, 2, \cdots, N\}$, and $ts_n^{(p)}$ is a realization of $TS^{(p)}$, $p \in \{1, 2, \cdots, P\}$. Let C be a discrete classification variable, with associate labels for each of the samples $\{c_1, c_2, \cdots, c_n\}$. C can take values in a finite set, but in this paper, only two class labels (1/0) are used.

2.1. Mutual information between time series

To estimate the mutual information between two time series $TS^{(p)}$ and $TS^{(q)}$, consider a set of "sliced" samples $\{\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_N\}$, where $\tilde{X}_n = (ts_n^{(p)}, ts_n^{(q)}), n \in \{1, 2, \cdots, N\}$ is a reduced sample of X_n with $ts_n^{(p)}$ and $ts_n^{(q)}$ being realizations of $TS^{(p)}$, and $TS^{(q)}$.

Let $\xi(n)$ be the distance from a sample $\tilde{X}_n$ to its kth nearest neighbor, where the distance $Dist(\tilde{X}_m, \tilde{X}_n)$ between two samples $X_m = (ts_m^{(p)}, ts_m^{(q)})$ and $X_n = (ts_n^{(p)}, ts_n^{(q)})$ are defined as the maximum value along the two variables

$$Dist(X_m, X_n) = max\{dist(ts_m^{(p)}, ts_n^{(p)}), dist(ts_m^{(q)}, ts_n^{(q)})\},$$

where $dist(ts^{(p)}, ts^{(q)})$ is the distance between two time series $ts^{(p)}$ and $ts^{(q)}$. Then one can count the number $\nu_{ts_n^{(p)}}$ of the time series $ts_m^{(p)}, m \in \{1, 2, \cdots, N\} - \{n\}$ whose distance $dist$ from $ts_n^{(p)}$ is equivalent or less than $\xi(n)$, i.e.,

$$\nu_{ts_n^{(p)}} = |\{m | m \in \{1, 2, \cdots, N\} - \{n\}, dist(ts_m^{(p)}, ts_n^{(p)}) < \xi(n)\}|.$$

Similarly, replace p by q, one can compute $\nu_{ts_n^{(q)}}$.

Following the method of Kraskov et al. (2004) and Ircio et al. (2020), the mutual information $I(TS^{(p)}; TS^{(q)})$ between two time series $TS^{(p)}$ and $TS^{(q)}$ is computed by

$$I(TS^{(p)}; TS^{(q)}) = \Psi(k) + \Psi(N) - \frac{1}{k} - \left(\frac{1}{N} \sum_{n=1}^{N} \left(\Psi(\nu_{ts_n^{(p)}}) + \Psi(\nu_{ts_n^{(q)}})\right)\right), \tag{1}$$

where $\Psi(x)$ is the digamma function, with $\Psi(x) = \Psi(x) + \frac{1}{x}$ and $\Psi(1) = -0.5772156$ (Euler-Mascheroni constant).

2.2. Mutual information between a time series and the classification variable

To estimate the mutual information between a time series TS and the classification variable C, consider the realizations $\{ts_1, ts_2, \cdots, ts_N\}$ of TS, and their associated class labels $\{c_1, c_2, \cdots, c_n\}$.

Let $d(n)$ be the distance from a time series sample ts_n to its kth nearest neighbors within the subset belong to c_n class. Then one can count the number ν_{ts_n} of time series $ts_m, m \in \{1, 2, \cdots, N\} - \{n\}$ whose distance from ts_n is equivalent or less than $d(n)$, i.e.,

$$\nu_{ts_n} = |\{m | m \in \{1, 2, \cdots, N\} - \{n\}, dist(ts_m, ts_n) < d(n)\}|.$$

Following the method of Ross (2014) and Ircio et al. (2020), the mutual information $I(TS; C)$

between the time series TS and the classification variable C is computed by

$$I(TS; C) = \Psi(k) + \Psi(N) - \Big(\frac{1}{N}\sum_{n=1}^{N}\big(\Psi(\nu_{c_n}) + \Psi(\nu_{ts_n})\big)\Big), \tag{2}$$

where $\Psi(x)$ is the digamma function as above and ν_{c_n} is the number of time series samples whose class labels are c_n.

2.3. Mutual information-based feature selection for time series classification

An optimal set of features should contain a list of highly informative and minimally redundant features. In our study, the minimum-Redundancy Maximum-Relevancy (mRMR) feature selection algorithm combined with mutual information is used to select the optimal set of time series features(Ircio et al., 2020). The goal of mRMR is to select a subset of features that maximize the score function

$$J(s) = \sum_{s \in S} I(TS^{(s)}; C) - \frac{1}{2|S|}\sum_{s \in S}\sum_{\substack{q \in S \\ q \neq s}} I(TS^{(s)}; TS^{(q)}), \tag{3}$$

where S is the index of the selected features. To achieve this goal, we apply a forward selection search strategy (Meyer et al., 2008): at the first step, the feature which has the largest mutual information with the class label, i.e., the one with the highest discriminating power, is selected; then at each of the following steps when the list of S' features are selected, mRMR ranks all the features $TS^{(q)}$ that are not selected according to $I(TS^{(s)}; TS^{(q)}) - \sum_{q \in S'} I(TS^{(s)}; TS^{(q)})$, and selects the top ranked one. In the meanwhile, at each step, the value of the score function is computed, and the set of features is finally determined when the score function is maximized.

3. Proposed experimental methodology

3.1. Data

Our experiment is based on the level-2 transaction data in the Chinese stock market. The stocks we consider are the constituent ones of China Security Index 300 (CSI 300) since they are the largest and most liquid stocks in the Chinese stock market and covers about 60% of the market's value. The period spans from January 2014 to August 2017 (a total of 896 trading days). The level-2 data is updated every 3 seconds, which consists of the current transaction price and volume, the cumulative number of trades and volume from the last record to the current one, as well as the best ten quotes just before the current transaction. The data are downloaded from Wind database(www.wind.com.cn). To ensure data adequacy, the stocks which are traded on less than 90% of the 896 trading days are deleted. To minimize the effect of extreme movement of stock prices on our analysis, the stocks of which the companies are newly listed, delisted, suspended in listing, or under special treatment are deleted. To this end, 189 stocks are retained for our analysis.

5

3.2. Intraday jump detection

To explore trading patterns before price jumps, the first key step is to detect accurately the jump components in the stock price series. Early studies use parametric models, such as GARCH-jump or SV-jump models, to estimate jumps, which always involve uncertain pre-setting of model forms and complicate parameter estimation(Maheu and McCurdy, 2004; Eraker et al., 2003). Recent development of jump detection methods in using high-frequency data introduce a type of non-parametric estimation of the jump component in asset prices, allowing more accurate identification of jumps at daily or even intraday levels (Barndorff-Nielsen and Shephard, 2006; Jiang and Oomen, 2008; Ait-Sahalia and Jacod, 2009; Corsi et al., 2010; Podolskij and Ziggel, 2010; Andersen et al., 2012; Lee and Mykland, 2008; Jiang and Zhu, 2017).

Our study concentrates on the intraday jumps. LM test, proposed by Lee and Mykland (2008), is the most widely used technique to detect intraday jumps. After a trading day is divided into M of Δt-minute intervals, the LM test uses the following jump detecting statistic to examine the presence of a jump in the ith interval of day t:

$$L_{t,i} = \frac{r_{t,i}}{\hat{\sigma}_{t,i}}, \tag{4}$$

where $r_{t,i}$ is the log return in this interval, and $\hat{\sigma}_{t,i}$ is the estimated instantaneous volatility, computed by the realized bipower variation of the returns in previous $K-2$ time intervals

$$\hat{\sigma}_{t,i}^2 = \frac{1}{K-2} \sum_{j=i-K+2}^{i-1} |r_{t,j-1}||r_{t,j}|. \tag{5}$$

Lee and Mykland (2008) derives the following rejection region for the null hypothesis at a significance level of α that no jump is present in this interval

$$\frac{L_{t,i} - C_{MT}}{S_{MT}} > -\log(-\log(1-\alpha)), \tag{6}$$

where
$$C_{MT} = \frac{\sqrt{2\log(MT)}}{c} - \frac{\log \pi + \log(\log(MT))}{2c\sqrt{2\log(MT)}}, S_{MT} = \frac{1}{c\sqrt{2\log(MT)}} \tag{7}$$

$c = \sqrt{2/\pi}$ and T is the total number of days.

The usage of a 5-minute interval to detect intraday jumps is a consensus choice in the existing literature(Bollerslev and Todorov, 2011; Liu et al., 2015; Wan et al., 2017), which represents a trade-off between maximizing statistical power and minimizing the effect of microstructure noise(Caporin et al., 2017). The Chinese Stock Exchange opens from 9:30 am to 11:30 am and 1:00 pm to 3:00 pm, with a total of 4 hours in a trading day. So a trading day is divided equivalently into 48 of 5-minute intervals. Then the LM method is used to detect the jumps within each interval. We

choose the previous interval window $K = 240$ for the detection following the suggestion of Lee and Mykland (2008) and Wan et al. (2017).

3.3. Candidate attributes to describe the micro-trading behaviors

Table 1 lists the candidate attributes we use to describe the micro-trading behaviors, which can be divided into two categories, liquidity measures and technical indicators. This set of attributes was initially used in our previous study, where they were processed into tabular data to fit into traditional machine learners for jump prediction(Kong et al., 2020). In this paper, we will use the time series of these attributes to study the micro-trading patterns. Formulas of these attributes can be found there and the computation of these attributes is summarized as follows.

We still base our computation on the sequence of 5-minute intervals in each trading day. The liquidity measures are simply computed for each stock within each interval, evaluating the market quality based on the level-2 information during that 5-minute period. The technical indicators are computed at the end of each interval using the 5-minute sampled price and volume information, describing the dynamics of the market trend in the near past. There are 18 technical indicators, while 12 of them requires the parameters of lagged periods. The lagged periods here are hence the number of lagged intervals, which are set to 5 and 20 to take account of the trends within a shorter and longer period. This actually gives us 30 technical indicators. To this end, the trading behaviors within a trading day is described by 40 attribute series of length 48, which are computed interval by interval of a day.

Table 1: Attributes

Liquidity measures		Technical indicators	
Return	r	Price rate of change	$PROC(q)$
Number of trades	K	Volume rate of change	$VROC(q)$
Trading volume	V	Moving average of price	$MA(q)$
Trading size	S	Exponential moving average of price	$EMA(q)$
Trade imbalance	TI	Bias to MA	$BIAS(q)$
Depth imbalance	DI	Bias to EMA	$EBIAS(q)$
Quoted spread	QS	Price oscillator to MA	$OSCP(q)$
Effective spread	ES	Price oscillator to EMA	$EOSCP(q)$
Realized volatility	RV	Fast stochastic %K	$fK(q)$
Cumulative return	R	Fast stochastic %D	$fD(q)$
		Slow stochastic %D	$sD(q)$
		Commodity channel index	$CCI(q)$
		Accumulation/Distribution oscillator	ADO
		True range	TR
		Price and volume trend	PVT
		On balance volume	OBV
		Negative volume index	NVI
		Positive volume index	PVI

In the technical indicators, q is the parameters of lagged period.

3.4. Attribute extraction

To explore the "abnormal" patterns of the attributes before price jumps, the main experiment of our study is to compare the dynamics of the attributes between the jumping and non-jumping samples. For the jumping samples, the attribute series are extracted within a 4-hour window before the occurrence of jumps. According to the literature, a 4-hour window is large enough to observe the abnormal patterns before price jumps(Boudt and Pertitjean, 2014; Wan et al., 2017). For the non-jumping samples, the time series are extracted within steady days. The steady days are defined as the days without jumps within the prior or subsequent 5 days.

It is noticed that some of the attributes have high idiosyncrasy (Podolskij and Ziggel, 2010; Boudt and Pertitjean, 2014). To compare across different stocks, days and intraday times, the liquidity measures need to be standardized; for a similar reason, we also standardize the technical indicators which has high idiosyncrasy. According to Boudt and Pertitjean (2014) and Kong et al. (2020), there are two standardization methods: one is performed by dividing each coordinate of the time series by the median of their data at the same time of the previous 60 days and subtracting 1; the other is performed by subtracting the median of their data at the same time of the previous 60 days from each coordinate of the time series. In our study, 15 types of attributes has high idiosyncrasy: U, K, S, RV, MA, EMA, TR, PVT and OBV are standardized by the first method, while OI, DI, QS, ES, $VROC$, NVI and PVI, which are already ratios, are standardized by the second one.

Besides, to avoid missing information in time series, we delete four types of samples in the dataset: First, we do not consider the samples in the first 60 days of each stock. Second, due to the 10% price change limit rule in the Chinese stock market, all the sequential limit-ups or limit-downs except the first one are deleted. Third, any sample within a day after the suspension of a stock is deleted. Forth, because of the implementation of a circuit breaker mechanism on the Chinese stock market from 2016/01/04 to 2016/01/07, during which the whole stock market halted several times, the samples between 2016/01/04 and 2016/01/08 for all stocks are deleted. Besides, dividends are subtracted from prices on dividend distribution dates, leading to a large change in prices, but they are not considered in our study either.

3.5. Stock representation

After the above data processing, we have obtained a group of time series for each attribute before the jumping times and on steady days respectively for each stock. To better reveal the statistical pattern of each stock, we then take the median values of the time series within each group for each stock. The median values are computed at each time coordinate. Figure 1 provides a general view of the stock representation methodology in this section.
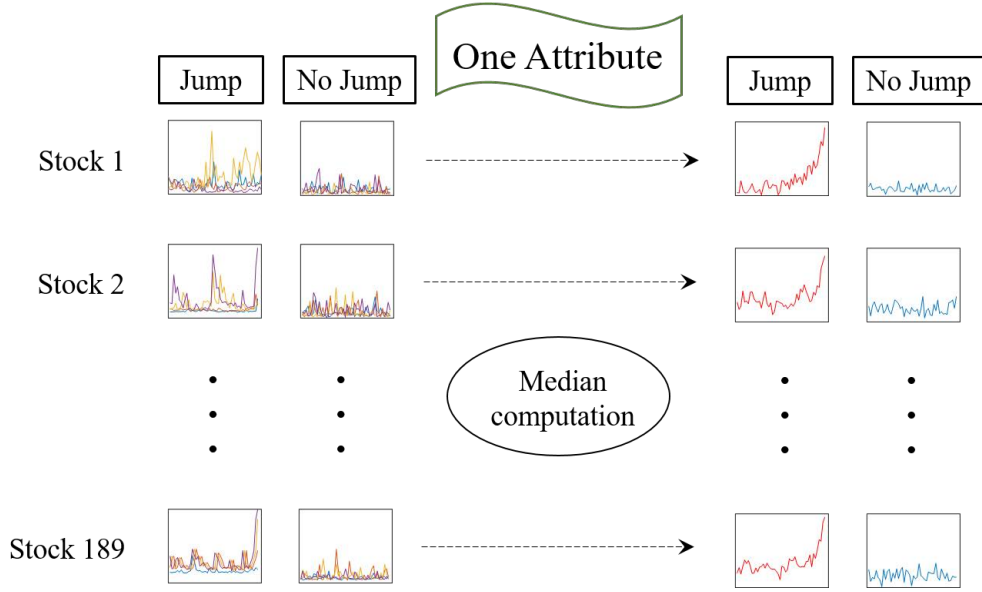
Figure 1: General view of the stock representation methodology

Prior to price jumps, we then have 189 time series respectively for each attribute. That is, before the occurrence of jumps, each stock is represented by a set of 40 time series. If we treat an attribute as a random variable, the 189 time series samples can be seen as its random realizations.

Similarly, on steady days, the median value of the time series for each attribute is computed for each stock. While the 4-hour information window under the jumping scenario can start from any time of a day, we should allow any starting point for the time series on steady days. Therefore, to perform a fair comparison, we randomize the order of the time series one by one on steady days and then compute their median values in each group. Then we have 40 time series to represent each stock on steady days. That is, we have 189 "virtual" time series respectively for each attribute, which can also be recognized as 189 realization samples of each attribute on steady days.

For the following analysis, we label the jumping and non-jumping samples as 1 and 0 respectively. To this end, we have 378 binary labeled samples $\{(X_1, C_1), (X_2, C_2), \cdots, (X_{378}, C_{378})\}$, where $C_i \in \{1, 0\}$ corresponds to the label of each sample, and $X_i = (ts_i^{(1)}, ts_i^{(2)}, \cdots, ts_i^{(40)})$, $i \in \{1, 2, \cdots, 378\}$ is represented by the 40 attribute time series of length 48.

3.6. Jump indicator selection

Using the method described in section 2.3, one can select a set of attributes, called jump indicators, which are highly related to the arrival of price jumps. The mutual information between each attribute and the label variable can be computed following section 2.2, while the mutual information between two time series is evaluated as outlined in section 2.1. In these computations, the distance $dist(ts^{(p)}, ts^{(q)})$ between two time series is needed. There are several types of distances

for time series in literature and we use three typical ones in our study, including the Euclidean distance, the Chebychev distance and the distance based on dynamic time warping.

Different from dynamic time warping, both the Euclidean distance and the Chebychev distance treat the time series as vectors. The Euclidean distance is simply the L2-norm distance between two vectors, while the Chebychev distance takes the maximum coordinate difference between the two vectors. Dynamic time warping(DTW), on the other hand, allows the comparison between two time series with varying lengths and speeds(Berndt and Clifford, 1994). In our study, although the attribute series are always extracted with the same length, their difference in the speed of change should be considered. In general, DTW searches for an optimal match between two vectors so that the Euclidean distances between their corresponding points is minimized. Besides, it has to comply with the following rules: first, every point in one sequence should be matched with one or more points of the other sequence; second, the first points from both sequence should be matched, and the same for the last points; third, the mapping of the points from one of the sequences to the other must be monotonically increasing, and vice versa. Detailed implementation of the DTW algorithm can be found in Berndt and Clifford (1994).

The estimation of mutual information based on the real data can be blurred by systematic errors resulting from finite-size issue(Steuer et al., 2002). To minimize this error, we correct the mutual information values by subtracting a zero-baseline. Random noise should ideally has zero mutual information with any random variable. So as in Steuer et al. (2002), we randomly permutate the realizations of the two variables, and compute the mutual information based on the surrogate pairs. This procedure is then repeated for many times (say 100) and the average of all the mutual information values are used as the zero-baseline.

3.7. Stock clustering according to jump indicator patterns

While the dynamics of the trading behaviors before price jumps is described by a set of jump indicators for each stock, one might wonder how the pattern differs from stock to stock and whether some of the stocks share similar patterns before price jumps. Clustering the stocks according to their jump indicator patterns can give a fast answer to this question, as well as to detect the stocks that have idiosyncratic trading patterns before price jumps.

Clustering is a task of partition the samples into several groups, where the samples in the same group are more similar to each other than to those in other groups. It is a widely used technique in pattern recognition and data mining problems. There are many types of clustering methods, such as connectivity-based clustering, centroid-based clustering, density-based clustering, and grid-based clustering. Different from other methods, connectivity-based clustering, also known as hierarchical clustering, do not provide a unique partitioning of the samples, but a dendrogram, showing how an extensive hierarchy of clusters merges with each other. The advantage of the method is that the

10

users can choose a set of appropriate clusters by choosing a cutoff of the inconsistency coefficient on the linkages of the dendrogram. Besides, one needs to choose the linkage criterion for computing the distance between clusters. In our study, the popular choice of "unweighted average linkage" is used.

In addition, clustering analysis can be easily dominated by the attribute with an extremely large scale. So to minimize the scale difference among jump indicators, each of the indicator series $(a_1, a_2, \cdots, a_T)$ are normalized before the clustering analysis by the min-max method

$$norma_t = \frac{a_t - minv}{maxv - minv},$$

where $minv$ and $maxv$ is the minimal and maximal value of a_t taken over all stocks.

4. Empirical results

4.1. Detected jumps

Following the method in section 3.2, we detect all the intraday jumps in the 5-minute sampled price series of the 189 sample stocks. To be more strict on the detected jumps, we set a significance level of 1%. Table 2 presents the total number of jumps counted over all the stocks as well as their average sizes. We can see that the number of positive jumps is larger than that of the negative jumps; while the average sizes of the negative jumps are larger. It can be explained by the fact that in the Chinese stock market most of the players are retail investors, who are more tending to chase after rising prices than falling ones given the condition that the market does not allow short sales; in the meanwhile, when prices are depreciating, the retail investors are more panic to short the stocks, resulting in larger negative jumps.

Table 2: Statistical description of detected jumps

	Positive jumps	Negative jumps
Number	20734	11597
Average return	0.0266	-0.0363

4.2. Dynamics of candidate attributes

To explore the abnormal trading behaviors of individual stocks before price jumps, we extract the attribute series within a 4-hour window, which is 48 intervals, before each intraday jump of each stock as outlined in section 3.4. Then prior to the occurrence of the jumps, each stock is represented by the median values of the attributes as outlined in section 3.5. To give a bird's eye view of the dynamics of all the attributes before price jumps, we plot the median value of the 189 samples before price jumps as in Boudt and Pertitjean (2014) and Wan et al. (2017)(Figure 2-3).

11

Figure 2: Dynamics of candidate attributes before price jumps. The solid line is the median values. The shaded region represents the range between 5% and 95% quantiles. The x-axis is the index of the intervals before price jumps.
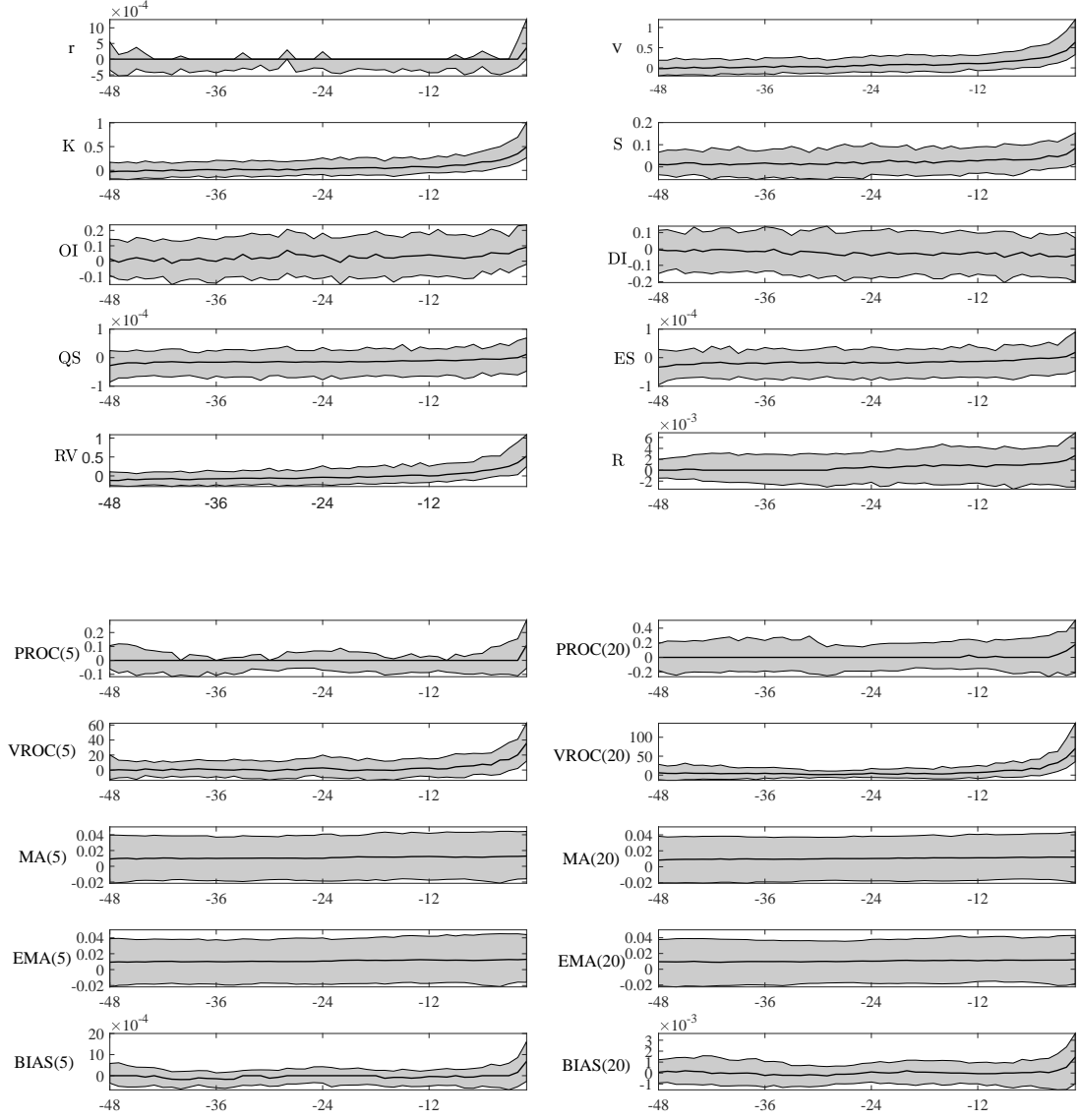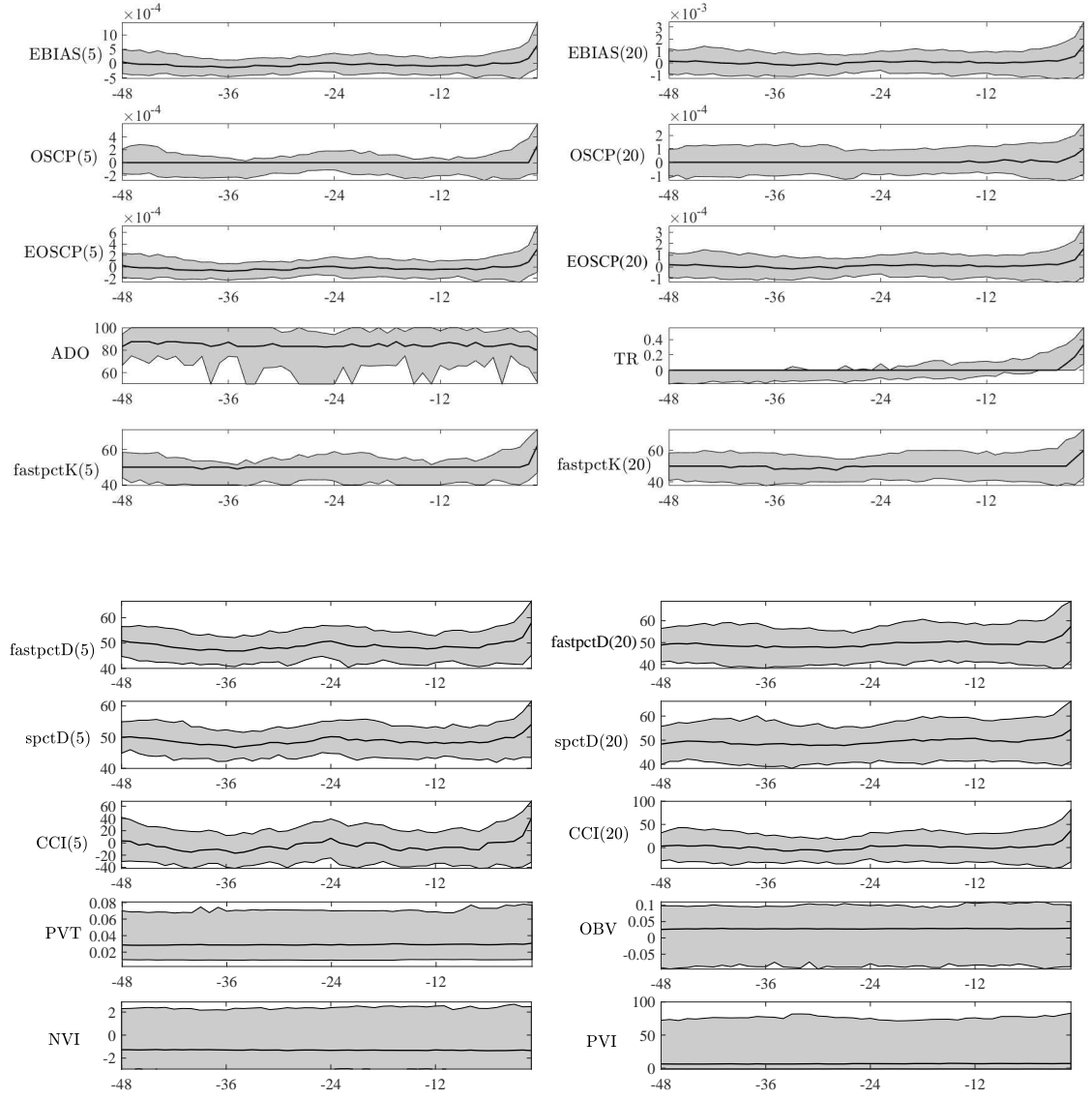
Figure 3: Dynamics of candidate attributes before price jumps. The solid line is the median values. The shaded region represents the range between 5% and 95% quantiles. The x-axis is the index of the intervals before price jumps.

From the figures, we find that a lot of the attributes, such as r, V, K, RV, $PROC$, $VROC$, $BIAS$, $OSCP$, TR, $fastpctK$, $fastpctD$, $spctD$, CCI, exhibit significantly abnormal surge just shortly before the occurrence of price jumps; while in most of the time, they maintain comparably stable. This might indicate that the abnormal movement only starts shortly before the occurrence of price jumps. The figures for the liquidity measures are very similar to those observed by Boudt and Pertitjean (2014) and Wan et al. (2017), who also compared these dynamics with the days without jumps through Mann-Whitney test. However, that conclusion is based on a point-by-point analysis, which does not take account of the characteristics of the whole time series; besides, it can not compare the informative levels among the attributes. We believe that further statistical analysis based on time series need to be conducted to explore the abnormal dynamics of the candidate attributes, for which we use the mutual information-based technique as follows.

4.3. Informative levels of candidate attributes

Mutual information between each attribute and the label variable, as defined in section 2.2, evaluates the levels of informative of each attribute with respect to the arrival of price jumps. Table 3 gives the mutual information values of the 40 attributes based on the three types of distances and 1, 3, or 5 nearest neighbors during the computation. To minimize the systematic errors resulting from finite-size issue, the zero-baseline has been subtracted as mentioned in section 3.6.

There is no meaning to compare the mutual information values but the relative rankings of the attributes across different distance or k parameter settings. Here a two-sided Wilcoxon signed rank test is adopted to compare the ranks of the attributes between any two of the nine scenarios. At the 1% significance level, all the tests fails to reject the null hypothesis of zero median in the rank difference. This indicates that the choice of the distance or the k parameters in general have low influence in comparing the information level of the attributes.

The boldfaced values in the table show that the trading volume V, the number of trades K, the realized volatility RV, the volume rate of change $PROC$ and the commodity channel index CCI have very high information levels with respect to the arrival of price jumps. V, K and $VROC$ are all volume-related attributes, the significant abnormality of which has also been observed in both Boudt and Pertitjean (2014) and Wan et al. (2017), indicating a demand for immediate execution before price jumps. RV and CCI are all volatility related measures, related to which Wan et al. (2017) observed that the volatility is significantly large before price jumps.

On the other hand, Boudt and Pertitjean (2014) and Wan et al. (2017) have also investigated the dynamics of other liquidity measures, such as QS, ES, OI, DI, before price jumps. According to the two studies, the abnormal movements of these measures are not very obvious. However, in our results, the information levels of OI and ES are moderate, while those of DI and QS are comparably lower; nevertheless the abnormality of trading behaviors embedded in these attributes

Table 3: Information level of candidate attributes

	Neighbors	Euclidean			Chebychev			DTW		
		1	3	5	1	3	5	1	3	5
1	r	0.36	0.32	0.30	0.32	0.28	0.24	0.23	0.21	0.21
2	V	**0.69**	**0.67**	**0.67**	**0.67**	**0.67**	**0.67**	**0.68**	**0.68**	**0.68**
3	K	**0.67**	**0.66**	**0.66**	**0.67**	**0.65**	**0.66**	**0.68**	**0.68**	**0.67**
4	S	0.49	0.49	0.48	0.47	0.47	0.47	0.53	0.51	0.51
5	OI	0.36	0.34	0.32	0.27	0.26	0.25	0.38	0.36	0.36
6	DI	0.22	0.22	0.21	0.17	0.16	0.17	0.26	0.25	0.25
7	QS	0.20	0.22	0.21	0.21	0.22	0.19	0.33	0.29	0.27
8	ES	0.29	0.26	0.25	0.23	0.24	0.23	0.41	0.39	0.36
9	RV	**0.67**	**0.67**	**0.67**	**0.67**	**0.66**	**0.64**	**0.68**	**0.68**	**0.66**
10	R	0.47	0.43	0.41	0.46	0.41	0.37	0.57	0.55	0.55
11	PROC(5)	0.51	0.50	0.38	0.47	0.38	0.37	0.48	0.41	0.41
12	PROC(20)	0.50	0.43	0.37	0.43	0.45	0.40	0.52	0.52	0.52
13	VROC(5)	**0.62**	**0.62**	**0.61**	**0.59**	**0.59**	**0.58**	**0.69**	**0.68**	**0.67**
14	VROC(20)	**0.66**	**0.68**	**0.67**	**0.68**	**0.67**	**0.67**	**0.67**	**0.66**	**0.67**
15	MA(5)	0.37	0.25	0.17	0.36	0.28	0.19	0.58	0.45	0.35
16	MA(20)	0.31	0.22	0.16	0.34	0.26	0.20	0.57	0.40	0.34
17	EMA(5)	0.35	0.25	0.18	0.33	0.26	0.22	0.57	0.44	0.36
18	EMA(20)	0.35	0.21	0.15	0.34	0.25	0.21	0.54	0.40	0.32
19	BIAS(5)	0.51	0.44	0.42	0.46	0.43	0.40	0.53	0.52	0.51
20	BIAS(20)	0.58	0.54	0.51	0.57	0.51	0.50	0.64	0.62	0.60
21	EBIAS(5)	0.56	0.52	0.51	0.46	0.47	0.47	0.57	0.56	0.54
22	EBIAS(20)	0.58	0.52	0.51	0.55	0.49	0.47	0.62	0.60	0.60
23	OSCP(5)	0.42	0.47	0.41	0.49	0.42	0.39	0.42	0.43	0.43
24	OSCP(20)	0.51	0.47	0.44	0.53	0.48	0.46	0.58	0.55	0.54
25	EOSCP(5)	0.56	0.52	0.50	0.46	0.47	0.48	0.56	0.57	0.55
26	EOSCP(20)	0.55	0.50	0.49	0.52	0.47	0.46	0.59	0.58	0.58
27	ADO	0.14	0.09	0.09	0.10	0.08	0.09	0.09	0.11	0.12
28	TR	0.60	0.57	0.52	0.60	0.58	0.56	0.58	0.59	0.59
29	fastpctK(5)	0.47	0.42	0.40	0.45	0.43	0.44	0.51	0.54	0.51
30	fastpctK(20)	0.47	0.47	0.42	0.53	0.45	0.44	0.62	0.61	0.60
31	fastpctD(5)	0.55	0.52	0.49	0.51	0.47	0.45	0.66	0.63	0.62
32	fastpctD(20)	0.55	0.53	0.48	0.51	0.47	0.45	0.65	0.65	0.64
33	spctD(5)	0.54	0.51	0.47	0.47	0.45	0.42	0.66	0.64	0.63
34	spctD(20)	0.45	0.42	0.41	0.44	0.45	0.43	0.67	0.65	0.64
35	CCI(5)	0.54	0.50	0.49	0.48	0.45	0.46	0.61	0.59	0.58
36	CCI(20)	**0.59**	**0.54**	**0.51**	**0.56**	**0.52**	**0.50**	**0.64**	**0.64**	**0.62**
37	PVT	0.22	0.12	0.09	0.21	0.16	0.13	0.31	0.18	0.13
38	OBV	0.15	0.09	0.04	0.16	0.07	0.06	0.40	0.20	0.12
39	NVI	0.14	0.08	0.05	0.13	0.09	0.06	0.28	0.19	0.13
40	PVI	0.19	0.09	0.04	0.24	0.14	0.10	0.36	0.21	0.10

The attributes ranked within the top 10 lists under all types of distances are boldfaced.

can not be ignored.

4.4. Informative window of candidate attributes

It seems from Figures 2 and 3 that the 4-hour window is too long to detect the abnormal dynamics of the candidate attributes, as most of the abnormality occurs just shortly before the arrival of price jumps. To examine the conclusion, further investigation based on time series analysis from the perspective of information theory is given as follows. Similar to Table 3, we compute the mutual information between each attribute with the class label variable, but shrink the prior 4-hour window(48 intervals) to 3 hours, 2 hours, 1 hour or 30 minutes long (i.e., 36, 24, 12 or 6 intervals). That is we evaluate the information level of each attribute within shorter periods right before price jumps.

There is no gold standard to choose the type of distance or the k parameter in k-nearest neighbors method (Ircio et al., 2020); but as verified in Table 3, the comparison among attributes does not alter very much with different choices of these parameters. To be concise, we only present the results in the main text based on the Euclidean distance and the 3-nearest neighbors method as in Table 4.

According to the definition of the mutual information in section 2.1 and 2.2, if the significant abnormality of an attribute occurs just shortly before the arrival of price jumps, ideally its information level within a shorter period of time should not differ much from those within the largest 4-hour window. One can observe from Table 4 that the information levels of most attributes, especially the highly informative ones, do not change very much as the prior window size shrinks from 48 intervals(4 hours) to 6 intervals(30 minutes). This indicates that a lot of the significant abnormalities indeed occur only shortly before the arrival of price jumps, which is consistent with the findings in existing literature (Boudt and Pertitjean, 2014; Wan et al., 2017; Będowska-Sójka, 2016). But there are still some attributes, which achieve much lower information levels within shorter prior windows, such as DI, ES, QS, RV, MA, EMA, $PROC$, TR. This implies that the abnormal movements of these attributes start much earlier and last for longer time until the occurrence of price jumps.

4.5. Jump indicators for pattern recognition

After evaluating the information levels of individual attributes, it is natural to consider the discriminating power of a combination of these attributes. Under the context of pattern recognition and machine learning, a combination of attributes tends to achieve better results. But it is important to check if there is high mutual dependency among the attributes. The mutual dependency can be evaluated by the mutual information outlined in section 2.1. Figure 4 shows the mutual dependency levels of pairs of candidates based on different distance and parameter settings. The

16

Table 4: Information level of candidate attributes in different sized prior windows

	Prior intervals	-48	-36	-24	-12	-6
1	r	0.32	0.34	0.33	0.29	0.33
2	V	0.67	0.67	0.65	0.65	0.62
3	K	0.66	0.66	0.64	0.64	0.61
4	S	0.49	0.48	0.49	0.49	0.52
5	OI	0.34	0.35	0.31	0.29	0.29
6	DI	0.22	0.22	0.19	0.15	0.11
7	QS	0.22	0.21	0.20	0.14	0.12
8	ES	0.26	0.24	0.22	0.19	0.17
9	RV	0.67	0.66	0.63	0.60	0.59
10	R	0.43	0.39	0.38	0.40	0.39
11	PROC(5)	0.50	0.50	0.45	0.43	0.36
12	PROC(20)	0.43	0.46	0.40	0.41	0.38
13	VROC(5)	0.62	0.63	0.64	0.65	0.65
14	VROC(20)	0.68	0.67	0.66	0.65	0.66
15	MA(5)	0.25	0.23	0.19	0.10	0.11
16	MA(20)	0.22	0.11	0.09	0.10	0.05
17	EMA(5)	0.25	0.22	0.18	0.15	0.13
18	EMA(20)	0.21	0.12	0.11	0.07	0.08
19	BIAS(5)	0.44	0.47	0.51	0.51	0.48
20	BIAS(20)	0.54	0.53	0.53	0.51	0.49
21	EBIAS(5)	0.52	0.53	0.56	0.52	0.49
22	EBIAS(20)	0.52	0.54	0.56	0.54	0.54
23	OSCP(5)	0.47	0.47	0.49	0.43	0.39
24	OSCP(20)	0.47	0.47	0.44	0.41	0.44
25	EOSCP(5)	0.52	0.54	0.55	0.51	0.51
26	EOSCP(20)	0.50	0.50	0.53	0.50	0.49
27	ADO	0.09	0.07	0.08	0.07	0.08
28	TR	0.57	0.56	0.56	0.49	0.48
29	fastpctK(5)	0.42	0.45	0.48	0.48	0.45
30	fastpctK(20)	0.47	0.46	0.47	0.48	0.45
31	fastpctD(5)	0.52	0.52	0.53	0.51	0.46
32	fastpctD(20)	0.53	0.52	0.49	0.45	0.44
33	spctD(5)	0.51	0.50	0.47	0.43	0.41
34	spctD(20)	0.42	0.40	0.37	0.38	0.40
35	CCI(5)	0.50	0.53	0.56	0.53	0.51
36	CCI(20)	0.54	0.51	0.52	0.50	0.46
37	PVT	0.12	0.12	0.11	0.11	0.12
38	OBV	0.09	0.07	0.06	0.03	0.06
39	NVI	0.08	0.09	0.07	0.03	0.01
40	PVI	0.09	0.08	0.10	0.04	0.02

These mutual information values are computed based on Euclidean distance and 3-nearest neighbors method.
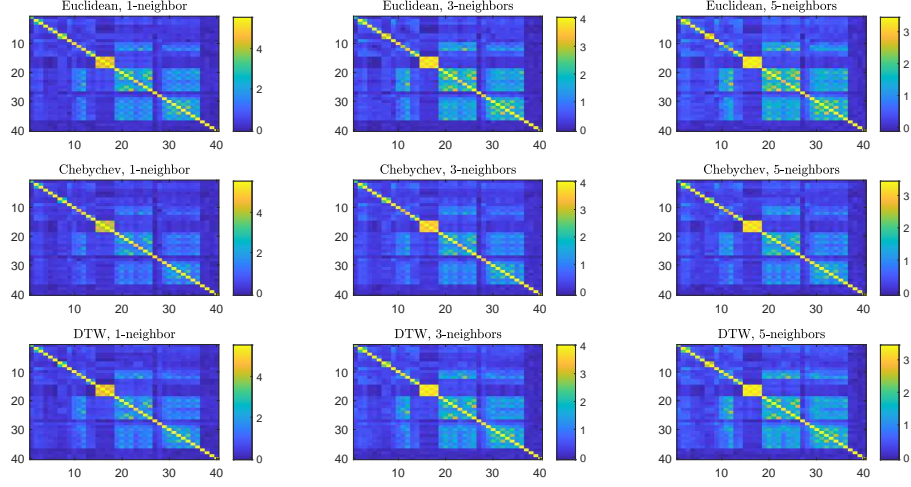
Figure 4: Mutual dependency level between pairs of candidates based on different distance and k-nearest neighbors method. Each block in the heatmaps represents the mutual dependency level of the two attributes whose index is denoted on the x- and y-axis.

nine heatmaps are very similar, and show that only a few of the attributes, such as MA's and EMA's are highly dependent, while most of the attributes have moderate or very low levels of mutual dependency.

The next task is to select a combination of attributes, called jump indicators, that are highly related to the arrival of price jumps. Because the abnormality of some attributes starts very early as concluded from Table 4, the feature selection technique, as presented in section 2.3, is performed on the set of attributes extracted within the prior 4-hour window. For comparison purposes, Table 5 exhibits the nine sets of selected jump indicators based on different types of distances and k values. It can be seen from the table that all sets of jump indicators include more than 70% of the candidate attributes, while the Chebychev and DTW distance sometimes even select almost all the jump attributes. There are up to 22 attributes that have been included in all the nine sets, which is in line with our previous conclusion that these attribute generally have low mutual dependency with each other. It is worthy to note that although some attributes have very low information level, such as ADO, OBV or PVI, and etc., they still play a role in the jump indicator set. Besides, even the MA's and EMA's are highly dependent, they are simultaneously selected in most cases, indicating that complementing information can still be found in these attributes.

4.6. Stock clustering according to jump indicator patterns

To cluster the stocks according to their trading abnormality before price jumps, the 189 jumping samples (that is the samples (X_i, C_i) with $C_i = 1$) as obtained in section 3.5 are adopted, each of which corresponds to a stock. To have a more robust clustering result with respect to different

18

Table 5: Selected jump indicators with different distances and k parameters

	Neighbors	Euclidean			Chebychev			DTW		
		1	3	5	1	3	5	1	3	5
1	r	√	√	√	√	√	√	√	√	√
2	V	√	√	√	√	√	√	√	√	√
3	K	√	√	√	√	√	√	√	√	√
4	S	√	√	√	√	√	√	√	√	√
5	OI	√	√	√	√	√	√	√	√	√
6	DI	√	√	√	√	√	√	√	√	√
7	QS			√	√	√	√	√	√	√
8	ES	√	√	√	√	√	√	√	√	√
9	RV	√	√	√	√	√	√	√	√	√
10	R	√	√	√	√	√	√	√	√	√
11	PROC(5)	√	√	√	√	√	√	√	√	√
12	PROC(20)	√				√	√			
13	VROC(5)	√	√	√	√	√	√	√	√	√
14	VROC(20)	√	√	√	√	√	√	√	√	√
15	MA(5)	√			√	√		√	√	√
16	MA(20)		√		√	√	√	√	√	√
17	EMA(5)	√	√	√			√	√	√	√
18	EMA(20)	√		√	√	√	√	√	√	√
19	BIAS(5)				√	√	√		√	√
20	BIAS(20)				√		√			√
21	EBIAS(5)			√		√	√	√	√	√
22	EBIAS(20)							√	√	√
23	OSCP(5)	√	√	√	√	√	√		√	√
24	OSCP(20)		√	√	√	√	√	√	√	√
25	EOSCP(5)		√				√		√	√
26	EOSCP(20)									
27	ADO	√	√	√	√	√	√	√	√	√
28	TR	√	√	√	√	√	√	√	√	√
29	fastpctK(5)	√	√	√	√	√	√	√	√	√
30	fastpctK(20)				√	√	√		√	√
31	fastpctD(5)				√	√	√	√	√	√
32	fastpctD(20)		√	√	√	√	√	√	√	√
33	spctD(5)	√	√	√	√	√	√	√	√	√
34	spctD(20)	√			√	√	√	√	√	√
35	CCI(5)	√	√	√	√	√	√	√	√	√
36	CCI(20)	√	√	√	√	√	√	√	√	√
37	PVT	√	√	√	√	√	√	√	√	√
38	OBV	√	√	√	√	√	√	√	√	√
39	NVI	√	√	√	√	√	√	√	√	√
40	PVI	√	√	√	√	√	√	√	√	√
	Total	28	28	29	33	35	37	33	37	38

The marker √ denotes the selected attributes in different cases.

distances, the 22 jump indicators which are covered by all the nine sets of Table 5 are used to represent each sample. Thus, each X_i is now represented by 22 time series of length 48. To minimize the effect of scale difference, each of the indicator series are normalized by the min-max method.

Figure 4.6 shows the hierarchical clustering results on the stocks based on the three types of distance. In the figures, the objects are linked by upside-down U-shaped lines. The height of the "U" represents the distance between the two clusters it links. If a link is approximately the same height as the links below it, it is said to exhibit a high level of consistency; in the contrary, it is said to be inconsistent with the links below it. In a cluster analysis, inconsistent links can provide the border of a natural division in a sample set. So from the following dendrogram, we can easily find the individual stocks that are very different from the majority.

Table 6: Most distinct stocks detected by the clustering analysis

Euclidean	Chebychev	DTW
19	157	19
20	164	20
157	186	157
164		164
177		177
186		188

In the figures, the most distinct stocks are includes in the cluster on the very right side of the trees. Comparing the height of the very top link to the below ones in each figure, we can find that the clustering analysis based on Chebychev distance recognizes a more different cluster to the majority than the other two types of clustering. Table 6 shows the index of the most distinct stocks detected by the three clustering analysis. Comparing to the Chebychev distance, the Euclidean and DTW distance detect larger and very similar sets of distinct stocks. Nevertheless, the stocks detected by the Chebycheve distance are included in the other two sets. It is likely that the Chebychev distance focuses on the maximal difference but ignores the small ones between time series; in this case, the Euclidean and DTW distance can capture more features of distinct stocks when performing the clustering, resulting in a larger set of distinct stocks.

Figure 4.6 gives, as an example, the attribute patterns of a stock (indexed by 157) before its price jumps, and shows how they are different from those of most stocks. For parsimony, we only present the dynamic patterns of the six attributes which have the highest discriminating power. In the eight plots, the solid lines represent the median attribute series of stock 157 before price jumps, while the shaded regions correspond to the range between 5% and 95% quantiles of the median attribute series of the stocks that are not present in Table 6. Comparing to most of the stocks, we notice that for this specific stock 157, the abnormal movements of the attributes V, K, $BIAS(20)$, $CCI(20)$, start much earlier, which is about 2 hours (24 intervals) before the

Hierarchical clustering based on Euclidean distance

Hierarchical clustering based on Chebychev distance

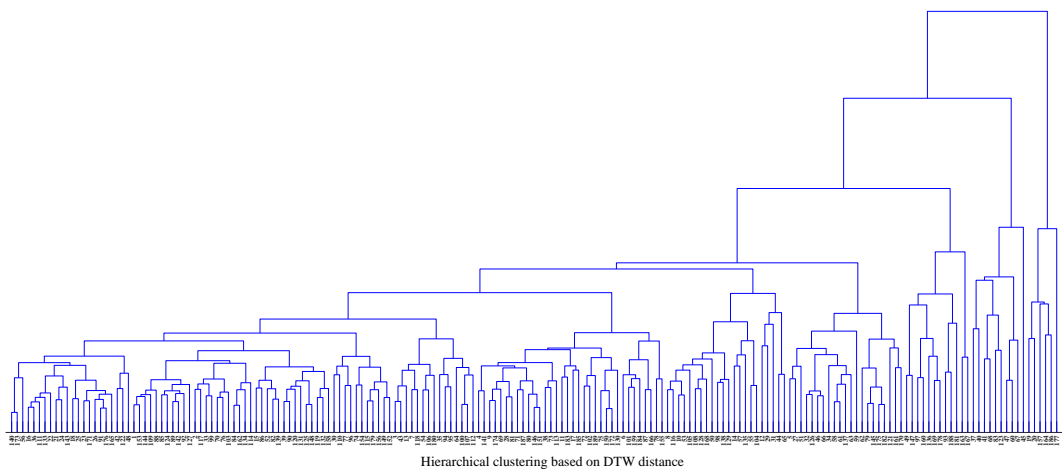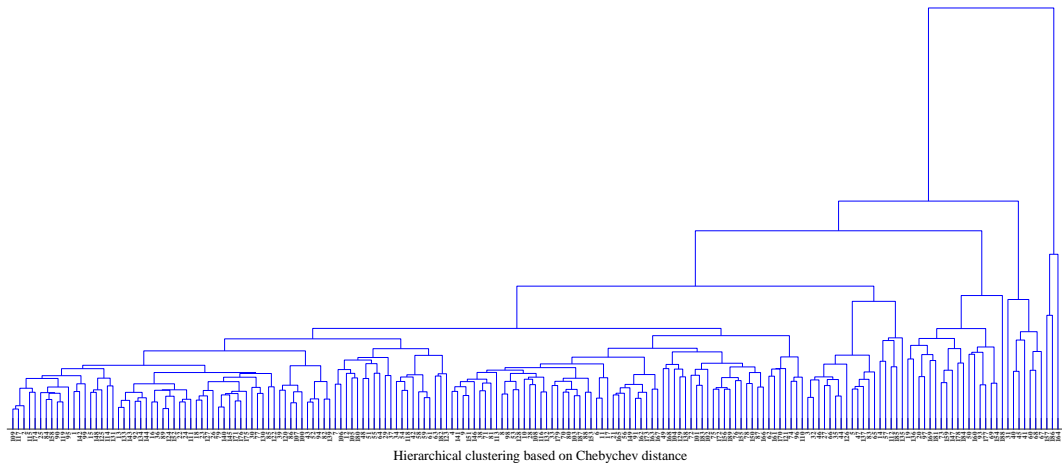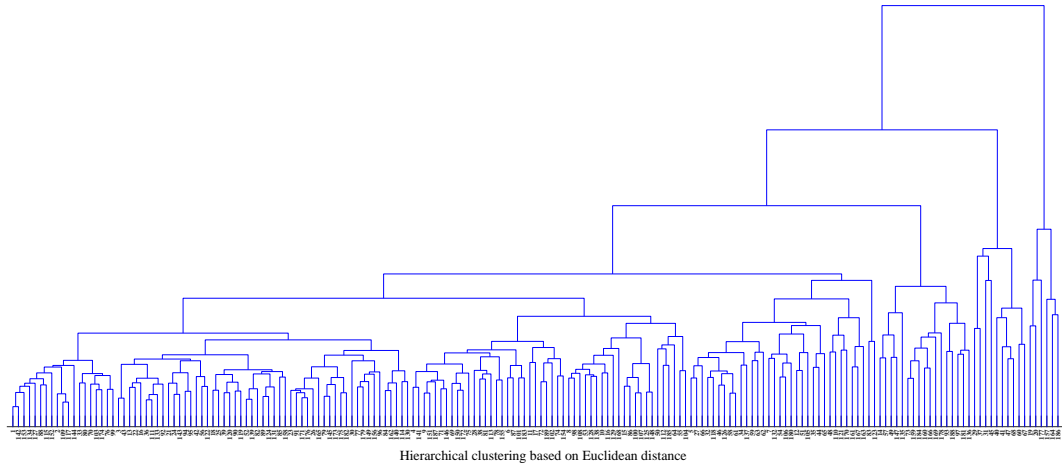Hierarchical clustering based on DTW distance

Figure 5: Hierarchical clustering of the stocks

occurrence of price jumps; meanwhile, their abnormality is very strong, which can be seen from the high levels of the solid lines equivalent or above the 5% quantile in their corresponding plots. On the other hand, the attributes RV, $VROC(5)$, $VROC(20)$ and TR predict the occurrence of the price jumps in a slower pace: they all reach very high levels but only within about a prior 0.5 hours (6 intervals). Nevertheless, the strong abnormality of these attributes have well explained why this stock is identified as a distinct stock by all the three clustering analysis from the perspective of micro-trading behavior before price jumps.
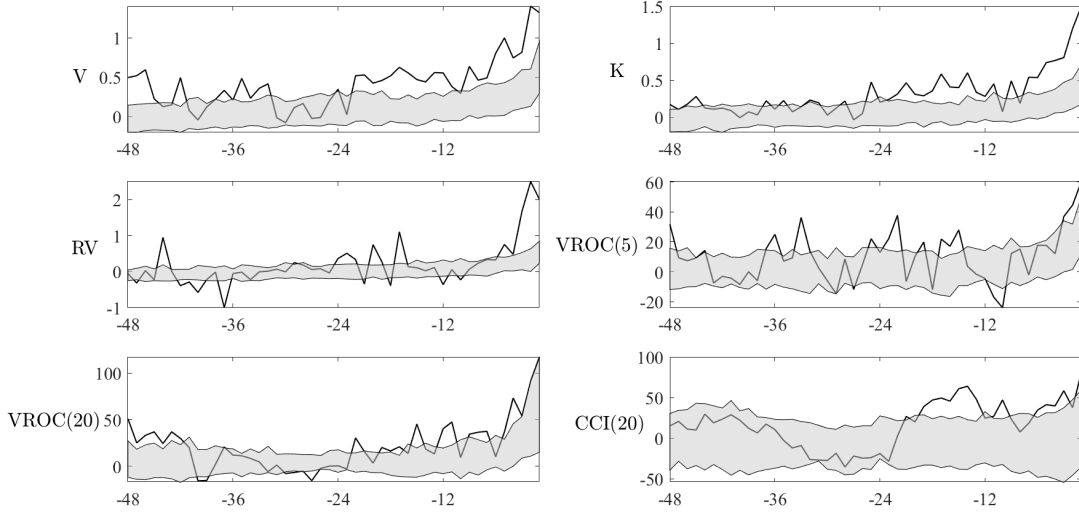
Figure 6: The trading patterns of a distinct stock (stock 157) before price jumps as an example. The solid line is the median time series of the corresponding attribute of the specific stock. The shaded region represents the range between 5% and 95% quantiles of the median time series of the stocks that are not present in Table 6. The x-axis is the index of the intervals before price jumps.

5. Discussion

Boudt and Pertitjean (2014) investigated the liquidity dynamics around price jumps in the U.S. stock market. Following that, Będowska-Sójka (2016) and Wan et al. (2017) performed similar experiments respectively on the Polish and Chinese stock markets. In these works, the Mann-Whitney tests on the jumping and non-jumping samples all showed some levels of abnormality in the liquidity measures related to the occurrence of price jumps. In this paper, we introduce a new method based on multivariate time series classification to better investigate the abnormal trading patterns before price jumps. This method can help to explore the "time-series" information in the candidate attributes used to describe the trading behaviors and to select a set of attributes, called jump indicators, to assist jump-related abnormal pattern recognition. In addition to the commonly

used liquidity measures, technical indicators are further included in the candidate attributes to describe the micro-trading behaviors from a different perspective.

Empirical study is conducted on the level-2 data of the constituent stocks of China Security Index 300 (CSI 300). After sample preparation, each stock is represented by a list of attribute series under two scenarios: before the intraday price jumps and on steady days. Evaluated by the mutual information, we found that some volume and volatility-related attributes, such as V, K, and RV, exhibit the most significant abnormality before price jumps, which is consistent with the findings in the existing literature(Boudt and Pertitjean, 2014; Wan et al., 2017). It is worthy to mention that the choice of distance and the k parameter in mutual information computation in general have low influence in comparing the information levels of different attributes. By evaluating their mutual information values within shrinking windows, we also find that even most of the attributes show abnormal movements just shortly before the occurrence of the price jumps, there are still some, such as QS, RV, MA, EMA, TR, start to move abnormally much earlier. Besides, the mutual information between time series shows that the attributes have low mutual dependency with each other. Based on that, we then select effectively a set of jump indicators and detect the stocks that have extremely abnormal trading behaviors before price jumps using clustering analysis. We believe that the methodology we proposed as well as the empirical results obtained in our study supplement the existing studies and provide a new perspective to investigate the micro-trading behaviors before price jumps in the financial market.

In future research, we will focus on the common and idiosyncratic micro-trading patterns of individual stocks and investigate the predictability of their price jumps using the jump indicators we selected. Based on that, high-frequency portfolio allocation and risk management strategies related to the occurrence of stock price jumps can also be designed.

References

References

Ait-Sahalia, Y., Jacod, J., 2009. Testing for jumps in a discretely observed process. The Annals of Statistic 37, 184–222.

Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169, 75–93.

Barndorff-Nielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4, 1–30.

Będowska-Sójka, B., 2016. Liquidity dynamics around jumps: The evidence from the Warsaw stock exchange. Emerging Markets Finance and Trade 52, 2740–2755.

Berndt, D., Clifford, J., 1994. Using dynamic time warping to find patterns in times series, in: AAAI-94 workshop on knowledge discovery in databases, pp. 359–370.

Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. The Journal of Finance 66, 2165–2211.

Boudt, K., Pertitjean, M., 2014. Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. Journal of Financial Markets 17, 121–149.

Caporin, M., Kolokolov, A., Renò, R., 2017. Systemic co-jumps. Journal of Financial Economics 126, 563–591.

Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics. Journal of Econometrics 159, 276–288.

Duffie, D., Pan, J., 2001. Analytical value-at-risk with jumps and credit risk. Finance and Stochastics 5, 155–180.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269–1300.

Fang, L., Zhao, H., Wang, P., Yu, M., Yan, J., Cheng, W., Chen, P., 2015. Feature selection method based on mutual information and class separability for dimension reduction in multidimensional time series for clinical data. Biomedical Signal Processing and Control 21, 82–89.

Farmer, J.D., Lillo, F., Mike, S., 2004. What really causes large price changes ? Quantitative Finance 4, 383–397.

Han, M., Ren, W., Liu, X., 2015. Joint mutual information-based input variable selection for multivariate time series modeling. Engineering Applications of Artificial Intelligence 37, 250–257.

He, G., Zhao, W., Xia, X., Peng, R., Wu, X., 2019. An ensemble of shapelet-based classifiers on inter-class and intra-class imbalanced multivariate time series at the early stage. Soft Computing 23, 6097–6114.

Ircio, J., Lojo, A., Mori, U., Lozano, J.A., 2020. Mutual information based feature subset selection in multivariate time series classification. Pattern Recognition 108, 107525.

Jarrow, R., Rosenfeld, E., 1984. Jump risks and the intertemporal capital asset pricing model. Journal of Business 57, 337–351.

Jiang, G., Oomen, R.C.A., 2008. Testing for jumps when asset prices are observed with noise–a "swap variance" approach. Journal of Econometrics 144, 352–370.

Jiang, G.J., Lo, I., 2011. Information shocks, liquidity shocks, jumps, and price discovery:Evidence from the U.S. treasury market. Journal of Financial and Quantitative Analysis 46, 527–551.

Jiang, G.J., Zhu, K.X., 2017. Information shocks and short-term market underreaction. Journal of Financial Economics 124, 43–64.

Joulin, A., Lefevre, A., Grunberg, D., Bouchaud, J.P., 2008. Stock price jumps: news and volume play a minor role. arXiv:0803.1769v1 .

Jovic, A., Jovic, F., 2017. Classification of cardiac arrhythmias based on alphabet entropy of heart rate variability time series. Biomedical Signal Processing and Control 31, 217–230.

Kapadia, N., Zekhnini, M., 2019. Do idiosyncratic jumps matter? Journal of Financial Economics 131, 666–692.

Kong, A., Zhu, H., Azencott, R., 2020. Predicting intraday jumps in stock prices using liquidity measures and technical indicators. Journal of Forecasting , Preprint.

Kraskov, A., Stögbauer, H., Grassberger, P., 2004. Estimating mutual information. Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics 69, 066138.

Lahaye, J., Laurent, S., Neely, C., 2011. Jumps, cojumps and maro announcements. Journal of Applied Econometrics 26, 893–921.

Lakhal, F., 2008. Stock market liquidity and information asymmetry around voluntary earnings disclosures New evidence from France. International Journal of Managerial Finance 4, 60–75.

Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563.

Liu, L., Patton, A.J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. Journal of Econometrics 187, 293–311.

Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics, and volatility components for individual stock returns. Journal of Finance 59, 755–793.

Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Meyer, P.E., Schretter, C., Bontempi, G., 2008. Information-theoretic feature selection in microarray data using variable complementarity. IEEE Journal of Selected Topics in Signal Processing 2, 261–274.

Podolskij, M., Ziggel, D., 2010. New tests for jumps in semimartingale models. Statistical Inference for Stochastic Processes 13, 15–41.

Ranaldo, A., 2008. Intraday market dynamics around public information arrivals intraday market dynamics around public information arrivals, in: Gregoriou, G.N., Lhabitant, F.S. (Eds.), Stock Market Liquidity: Implications for Market Microstructure and Asset Pricing. John Wiley and Sons, Hoboken(NJ), pp. 199–226.

Riordan, R., Storkenmaier, A., Wagener, M., Zhang, S.S., 2013. Public information arrival : Price discovery and liquidity in electronic limit order markets. Journal of Banking and Finance 37, 1148–1159.

Ross, B.C., 2014. Mutual information between discrete and continuous data sets. PLoS ONE 9.

Shannon, C.E., 2001. A mathematical theory of communication. ACM SIGMOBILE Mobile computing and Communication Review 5, 3–55.

Steuer, R., Kurths, J., Daub, C.O., Weise, J., Selbig, J., 2002. The mutual information: Detecting and evaluating dependencies between variables. Bioinformatics 18, 231–240.

Wan, D., Wei, X., Yang, X., 2017. Liquidity dynamics around intraday price jumps in Chinese stock market. Journal of Systems Science & Complexity 30, 434–463.

Yoon, H., Yang, K., Shahabi, C., 2005. Feature subset selection and feature ranking for multivariate time series. IEEE Transactions on Knowledge and Data Engineering 17, 1186–1198.

Zhou, C., Wu, C., Wang, Y., 2019. Dynamic portfolio allocation with time-varying jump risk. Journal of Empirical Finance 50, 113–124.